

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-322044

(43)Date of publication of application : 12.12.1997

(51)Int.Cl.

H04N 5/225  
G10L 3/00  
H04N 7/15  
H04R 1/02

(21)Application number : 08-137820

(71)Applicant : SHARP CORP

(22)Date of filing : 31.05.1996

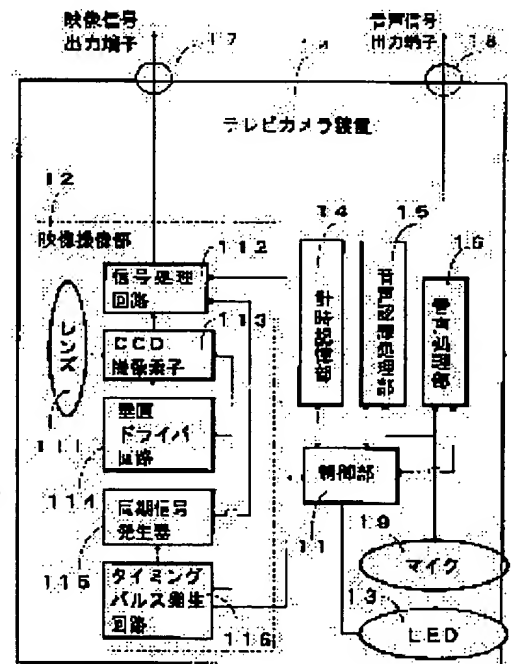
(72)Inventor : YAMAMOTO MAKOTO

## (54) TELEVISION CAMERA DEVICE

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To stop transmission or transmit holding music through a computer system or the like instead of conversation voice and secure privacy in the case that sound recording desired to be temporarily evaded is present by providing a shifting means for shifting to a control mode for not sending out voice gathered by a microphone to the outside of a television camera device by light-shielding the front surface of an image pickup lens for prescribed time.

**SOLUTION:** The control part 11 of this television camera device 1a controls the whole body and a video image pickup part 12 picks up the video images of an object, converts them to television video signals and outputs them. A timing storage part 14 stores the set value of a timer operation and a voice recognition processing part 15 identifies a vocabulary determined beforehand from gathered voice and converts it to a control code. A voice processing part 16 performs a processing to voice signals like interrupting voice signals from the microphone 19 and sending out the holding music or the like instead thereof.



## LEGAL STATUS

[Date of request for examination]

21.01.2000

[Date of sending the examiner's decision of rejection]

**\* NOTICES \***

**Japan Patent Office is not responsible for any damages caused by the use of this translation.**

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

**DETAILED DESCRIPTION**

---

**[Detailed Description of the Invention]**

[0001]

[Field of the Invention] This invention relates to the television conference which transmits the image of the television camera equipment which picturizes an image to a remote place using means of communications, and the television camera equipment which can process into arbitration voice which used it for TV phone equipment and collected the sound especially about suitable television camera equipment.

[0002]

[Description of the Prior Art] The TV phone or video conference system which also used image information as a mean to plan KOMYUNYUKESHON with those who are present in the conventionally distant location in addition to speech information is known.

[0003] The technique perform with voice control action of the television camera equipment used for the above-mentioned TV phone or a video conference system is already known, and the example which separates control voice from the conversation which is keynote voice using two microphones apart from the technique performed by change actuation of a switch as a means to separate control voice from the conversation which is keynote voice is indicated by the publication-number No. 61497 [ five to ] official report. Moreover, the example which separates control voice from the conversation which is keynote voice by performing frequency analysis for the voice which collected the sound is indicated by the publication-number No. 173592 [ five to ] official report.

[0004]

[Problem(s) to be Solved by the Invention] However, although it will be necessary to protect one's privacy as a result of an opportunity to mix an image and voice and communicate with many and unspecified partners becoming possible with the advance of a communication network, [, such as an image, voice, etc. not to open to a partner, ] Control of the television camera equipment by the conventional keyboard, a mouse, etc. had complicated actuation, and its configuration of a device was very complicated in the technique of dividing control voice into the conversation going-up automatic target which is keynote voice.

[0005] It is shifting to the control mode only in a motion of a hand, without being dependent on the class of computer apparatus which this invention is originated in view of such a situation, and is connected. While realizing easily sending-out cutoff of conversation voice not to open to a partner Since only recognition of the control voice of only the specific vocabulary defined beforehand, without being influenced in the control mode by the conversation which is keynote voice is performed, simplification of a speech recognition device can be attained and it is in offering the television camera equipment which can realize processing of the sending-out voice to a partner easily by the voice control.

[0006]

[Means for Solving the Problem] In order to solve the above-mentioned technical problem, it connects with the system which uses a communication line and transmits an image, and the next configuration is used for this invention in television camera equipment equipped with an image pick-up means to picturize an image, and a sound-collecting means to collect voice with a microphone.

[0007] Namely, the television camera equipment concerning invention according to claim 1 By equipping a control section with the function to hold the hysteresis of the electrical signal level from the CCD image sensor always measured in order to perform white balance control for guaranteeing the color reproduction nature of a color image By

acquiring the information whose period when the electrical signal level from a CCD image sensor decreased rapidly from the hysteresis is predetermined that time amount continuation was carried out, it can know that the hand shaded the front face of an image pick-up lens.

[0008] Moreover, the television camera equipment concerning invention according to claim 2 the speech recognition means of the speech information obtained with a sound-collecting means, a speech processing means to perform predetermined processing to the voice sent out in the exterior of television camera equipment, and a time check -- by having a storage means the time amount progress while stopping sending out in the exterior of the control voice which collected the sound with the microphone in the control mode, after shifting to the control mode -- a time check -- termination of the control mode can be detected because it is in agreement with the setting information on a storage means, and television camera equipment can be automatically returned to normal mode actuation.

[0009] And the television camera equipment concerning invention according to claim 3 comes to have the control means which controls actuation of television camera equipment by the speech recognition means of the control voice which consists of vocabularies of the specification defined beforehand which collected the sound with the microphone in a configuration according to claim 2.

[0010] And since the display action according to the sending-out existence to the partner of the voice which came to have the display means which can display lighting or a putting-out-lights condition on arbitration by control [ control section ], therefore was collected with the microphone, the recognition propriety of control voice, and \*\*\*\*\* of the control mode is performed, the television camera equipment concerning invention according to claim 4 can know in what kind of operating state television camera equipment is [ a user ] again.

[0011]

[Embodiment of the Invention] Hereafter, the gestalt of operation of the television camera equipment concerning this invention is explained to a detail based on a drawing. Drawing 1 is the block diagram showing the outline of a video conference system in which the television camera equipment of this invention was carried out.

[0012] In drawing 1, after the image and voice which were picturized with one television camera equipment 1a are transmitted to computer apparatus 2a connected by the video-signal line and the public telephone sound signal line, they are told to the display and the sound-reinforcement section of computer apparatus 2b in the location distant via the communication networks 3, such as a circuit.

[0013] Similarly, after the image and voice which were picturized with the television camera equipment 1b of another side are transmitted to computer apparatus 2b connected by the video-signal line and the sound signal line, they are told to the display and the sound-reinforcement section of computer apparatus 2a in the location distant via the communication networks 3, such as a dial-up line.

[0014] Drawing 2 is the block diagram showing the whole television camera equipments [ 1 ] and 1b configuration of drawing 1, and drawing 3 is the perspective view showing the appearance of the television camera equipments 1a and 1b of drawing 1.

[0015] These television camera equipments 1a and 1b Picturize the image of the control section 11 which controls the whole equipment, and a photographic subject, and it changes into a television video signal. the time check which memorizes the set point of the image image pick-up section 12 to output and timer actuation -- the sound signal from the storage section 14, the speech recognition processing section 15 which changes a specific vocabulary into a control sign from the inside of the collected voice, and a microphone being intercepted, or It has the speech processing section 16 which processes to a sound signal, such as sending out tone-on-hold comfort instead.

[0016] And while the lens 111 for an image pick-up is formed in the front face of the television camera equipments 1a and 1b, by the microphone 19 which collects voice, and the control section 11, from the case 24, LED13 by which lighting control is carried out is horizontally [ the ] exposed, and is arranged. Moreover, the electric power switch 20 is arranged at the flank of television camera equipment.

[0017] Moreover, the control section 11 has an internal timer and nonvolatile memory (all are illustration abbreviations), and while clocking always exact time of day because an internal timer carries out counting of the frequency of an oscillator circuit, the specific vocabulary which should be recognized as a control sign in the digital sign and the speech recognition processing section 15 of the note which can be recognized as music is beforehand memorized by changing into nonvolatile memory by the D/A conversion circuit of the speech processing section 16.

[0018] Furthermore, the image pick-up section 12 Light The image of the CCD image sensor 113 and photographic subject which are changed into an electrical signal The lens 111 for making the front face of CCD image sensor 113

carry out image formation optically, a CCD image pick-up -- the electrical signal from bamboo The synchronizing signal generator 115 synchronous-related [ for changing into the digital disposal circuit 112 and television video signal which are changed into the PAL television video signal constituted from an image of 50 sheet \*\* in the NTSC television video signal constituted from an image of 60 sheets in 1 second, or 1 second ], and a CCD image sensor It consists of the timing pulse generating circuits 116 and the perpendicular driver circuits 114 for driving.

[0019] and the above-mentioned image pick-up section 12 -- the image pick-up means of a claim -- the image processing section 13 -- the image processing means of a claim -- the speech recognition processing section 15 support to the speech recognition means of a claim, and a control section 11 and LED13 support [ the speech processing section 16 ] the speech processing means of a claim at the display means of a claim, respectively.

[0020] Next, actuation of the television camera equipments 1a and 1b of the above-mentioned configuration is explained with reference to the state transition diagram shown in drawing 4 . For example, if the electric power switch section 20 of one television camera equipment 1a is operated and a power source is switched on, television camera equipment 1a will be in a normal operation condition, and the television video signal from the image image pick-up section 12 will be transmitted to KOMBYUTA equipment 2a from the image output terminal 17.

[0021] After the volume of the conversation voice similarly collected with the microphone 19 is controlled in the speech processing section 16 based on the amplification degree set up by the control section 11, it is transmitted to KOMBYUTA equipment 2a from the voice output terminal 18. Moreover, LED13 which is the display device of television camera equipment 1a is in the condition of having always switched on the light, and it has reported to the use that television camera equipment is operating in normal operation mode. Therefore, while the image of a photographic subject is displayed on the display of computer apparatus 2a, it is in the condition that conversation voice is heard from the loudspeaker.

[0022] When a reason to perform conversation which be not asked to the user by the side of computer apparatus 2b in this normal operation mode occurs, the user of television camera equipment 1a Without changing wrap habit for the speaker of the conventional telephone by hand, if wrap actuation is performed 1 second or more, the microphone section of television camera equipment 1a Since it approaches and the lens 111 currently installed is also covered by hand, an image input will be in a protection-from-light condition. In order that a control section 11 may acquire the hysteresis that the period when the electrical signal level from CCD image sensor 113 shifted to black level rapidly continued for 1 second, through a digital disposal circuit 112, a control section shifts to the control mode while performing cutoff directions in the speech processing section 16 immediately.

[0023] In the control mode, for amplification degree 0, since the conversation voice collected with the microphone 19 since voice cutoff directions had appeared in the speech processing section 16 from the control section 11 cannot be transmitted to computer apparatus 2a and transmitted from the voice output terminal 18 to computer apparatus 2b which is naturally a partner, either, it can secure privacy. Moreover, LED13 which is the display device of television camera equipment 1a is shifting to the condition of having always put out the light, and it has reported to the user that television camera equipment 1a is operating in control action mode.

[0024] Since the sound-collecting actuation with a microphone 19 is continued also in the control mode and, as for this collected voice, speech recognition actuation is performed in the speech recognition processing section 15 in the condition of the control mode, When a user wants to change actuation of television camera equipment In order to avoid incorrect recognition, after uttering first the word "control", predetermined modification for every control word of the i performed to television camera equipment by uttering the control word which consists of a vocabulary which was defined beforehand continuously and set as the control section 11:

[0025] the case where the speech recognition processing section 15 has recognized a user's control voice at this time -- LED13 -- the period of 1 second -- a 2 times point -- it is the configuration that carrying out decrease actuation reports a recognition situation to a user. Here, the incorrect recognition rate of the speech recognition processing section 15 is "small" lowered "greatly" by the "tone-on-hold comfort" shown in drawing 4 , silent [ "silent" ], although it is the control voice by which speech recognition is carried out in the control mode, and limiting to six kinds of "X second" and "a return", and positive recognition is enabled.

[0026] The contents of control of the television camera equipment for every control word of this are explained below. First, when "tone-on-hold comfort" has been recognized as a control word, tone-on-hold comfort is generated because the control section 11 of television camera equipment changes the digital sign of the note beforehand stored in the nonvolatile memory with which the interior was equipped by the D/A conversion circuit of the speech processing

section 16, and the tone-on-hold comfort is always sent out from the voice output terminal 18 between the control modes.

[0027] Moreover, when "it is silent" has been recognized as a control word, the control section 11 of television camera equipment is [ that no voice is sent out ] silent from the voice output terminal 18, in order to stop the sending out, where tone-on-hold comfort is sent out from the speech processing section 16. Naturally, in the control mode, it is immediately after shift and, in the case of a silent sending-out condition, no change is produced.

[0028] and -- as a control word -- "-- small -- " -- when it has recognized, the control section 11 of television camera equipment changes a setup so that it may become small to the conventional setup about the amplification degree at the time of amplifying the conversation voice which collected the sound with the microphone 19 after returning to normal operation mode in the speech processing section 16, and sending out from the voice output terminal 18.

[0029] and -- again -- as a control word -- "-- large -- " -- when it has recognized, the control section 11 of television camera equipment changes a setup so that it may become large to the conventional setup about the amplification degree at the time of amplifying the conversation voice which collected the sound with the microphone 19 after returning to normal operation mode in the speech processing section 16, and sending out from the voice output terminal 18.

[0030] furthermore, when "X second" has been recognized as a control word, the control section 11 of television camera equipment has memorized the setup time of \*\* which carries out a normal operation MODOHE auto return -- a time check -- when the setup time of the storage section 14 has been recognized to be 10 seconds as a control word, it is changed at 10 seconds.

[0031] the elapsed time when "the return" has been recognized as a control word, after the control section 11 of television camera equipment shifts further again from the normal operation mode clocked by the internal timer -- a time check -- even if it is in the condition of having not reached the setup time of the storage section 14, it shifts to LED flashing mode immediately.

[0032] although explanation of the contents of control of the television camera equipment for every control word is above, in control action mode, the elapsed time after shifting from normal operation mode always clocks by the internal timer of a control section 11 -- having -- \*\*\*\* -- a time check -- it shifts to LED flashing mode because it is in agreement with the setup time of the storage section 14.

[0033] While reporting that it is near for LED13 to perform question flashing actuation for 5 seconds in a cycle of 1 seconds in LED flashing mode, and to shift to normal operation mode at a user Voice output cutoff is canceled for 5 seconds after progress, and after the volume of the conversation voice collected with the microphone 19 is controlled in the speech processing section 16 based on the amplification degree set up by the control section 11, it returns to the normal operation mode transmitted to KOMBYUTA equipment 2a from the voice output terminal 18.

[0034] If it returns to normal operation mode, LED13 which is the display device of television camera equipment 1a will be always switching on the light, and it will report to a user that television camera equipment is operating in normal operation mode. Therefore, while the image of a photographic subject is displayed on the display of computer apparatus 2a, it returns to the condition that conversation voice is also heard from the loudspeaker.

[0035] Henceforth, although the return actuation to the control mode and normal operation mode is repeated whenever a lens 111 is shaded more than for 1 second, when an electric power switch 20 is again switched on since the energization to television camera equipment is intercepted and a control section is reset even if it intercepts an electric power switch 20 in the state of the control mode with a natural thing, it will operate in normal operation mode.

[0036]

[Effect of the Invention] Since the television camera equipment of this invention is the above configurations, since invention according to claim 1 changes a halt or tone-on-hold comfort of transmission into conversation voice and can transmit it through a computer apparatus etc. to avoid temporarily that the situation which be not known as a partner arises and the present conversation situation is collected with television camera equipment, it can secure privacy.

[0037] Moreover, invention according to claim 2 can offer the control means of the operating state of television camera equipment easily by establishing the control mode which performs return actuation automatically easily by easy actuation of shading the lens section of television camera equipment that it can shift.

[0038] And in the case of a control vocabulary, the system which performs control action of television camera \*\*\*\* with voice can consist of easily performing speech recognition of the control voice which should limit invention according to claim 3 only to the voice collected in the control mode, and should be recognized as a control sign, such as prefacing language by means of which conversation voice and discernment become easy, without being influenced by

conversation voice.

[0039] And the check of whether the voice collected with the microphone is sent out to the partner is not only easy, but invention according to claim 4 can enable the check of the return to the normal mode being as Hasama from the recognition situation of the control voice in the control mode, and the control mode easily visually again.

---

[Translation done.]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-322044

(43) 公開日 平成9年(1997)12月12日

| (51) Int.Cl. <sup>6</sup> | 識別記号  | 序内整理番号 | F I           | 技術表示箇所  |
|---------------------------|-------|--------|---------------|---------|
| H 0 4 N 5/225             |       |        | H 0 4 N 5/225 | Z       |
| G 1 0 L 3/00              | 5 5 1 |        | G 1 0 L 3/00  | 5 5 1 G |
| H 0 4 N 7/15              |       |        | H 0 4 N 7/15  |         |
| H 0 4 R 1/02              | 1 0 7 |        | H 0 4 R 1/02  | 1 0 7   |

審査請求 未請求 請求項の数 4 O L (全 6 頁)

(21) 出願番号 特願平8-137820

(22) 出願日 平成8年(1996)5月31日

(71) 出願人 000005049

シャープ株式会社

大阪府大阪市阿倍野区長池町22番22号

(72) 発明者 山本 誠

大阪府大阪市阿倍野区長池町22番22号 シ

ャープ株式会社内

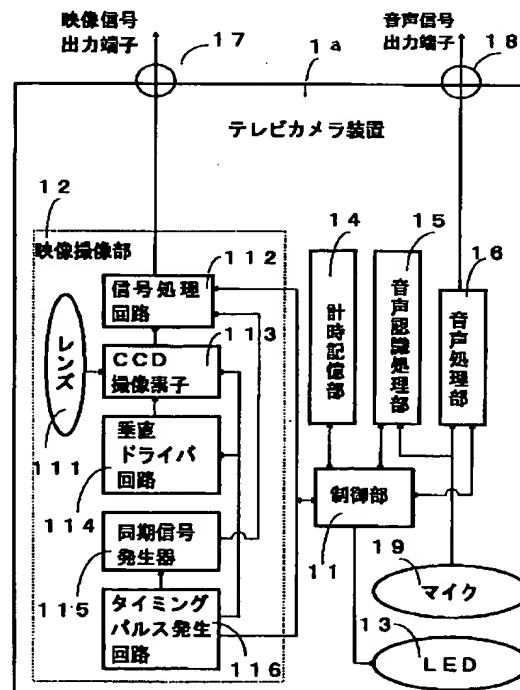
(74) 代理人 弁理士 梅田 勝

(54) 【発明の名称】 テレビカメラ装置

(57) 【要約】

【課題】 相手に公開したくない映像、音声等自分のプライバシーを守る必要が生じているが、従来のキーボード、マウス等によるテレビカメラ装置の制御は操作が複雑であり、また、主音声である会話上り自動的に制御音声を分離する手法においては、機器の構成が極めて複雑であった。

【解決手段】 撮像レンズの前面を所定の時間遮光することによって、マイクで集音した音声をテレビカメラ装置1 a、1 bの外部へ送出しない制御モードに移行する移行手段を有してなるものである。



## 【特許請求の範囲】

【請求項1】 通信回線を用いて映像を伝送するシステムに接続されるものであって、映像を撮像する撮像手段と、マイクで音声を集音する集音手段とを備えたテレビカメラ装置において、

上記撮像レンズの前面を所定の時間遮光することによって、マイクで集音した音声をテレビカメラ装置の外部へ送出しない制御モードに移行する移行手段を有することを特徴とするテレビカメラ装置。

【請求項2】 通信回線を用いて映像を伝送するシステムに接続されるものであって、映像を撮像する撮像手段と、マイクで音声を集音する集音手段と撮像レンズの前面の遮光を検出する信号処理手段とを備えたテレビカメラ装置において、

撮像レンズの前面の遮光を検出することにより移行した制御モードにて、上記集音手段で得られる音声情報の中よりあらかじめ定められた語彙を制御符号として識別を行う音声認識手段と、テレビカメラ装置の外部へ送出される音声に所定の処理を行う音声処理手段と、制御モードより通常モードへ復帰する所定の時間経過管理を行う計時記憶手段とを有することを特徴とするテレビカメラ装置。

【請求項3】 通信回線を用いて映像を伝送するシステムに接続されるものであって、映像を撮像する撮像手段と、マイクで音声を集音する集音手段と、上記集音手段で得られる音声情報の中よりあらかじめ定められた語彙を制御符号として識別を行う音声認識手段とを備えたテレビカメラ装置において、

上記集音手段で得られる会話音声の音声信号出力端子より装置外部への送出レベルの制御、音声信号出力端子より装置外部への送出音声として無音若しくは保留音楽の選択切り替えの制御、あるいは計時記憶手段の設定内容の変更制御動作が、制御モード時に音声認識手段で識別した制御符号により制御する制御手段を有してなることを特徴とするテレビカメラ装置。

【請求項4】 請求項1記載のテレビカメラ装置において、

マイクで集音した音声を会話音声としてテレビカメラ装置の外部へ送出する通常モードと、マイクで集音した音声を制御音声とするためテレビカメラ装置の外部へ送出しない制御モードの識別の表示手段とを有することを特徴とするテレビカメラ装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、映像を撮像するテレビカメラ装置の映像を通信手段を用いて遠隔地へ伝送するテレビ会議、テレビ電話装置に使用して好適なテレビカメラ装置に関し、特に集音した音声の加工を任意に行い得るテレビカメラ装置に関するものである。

## 【0002】

【従来の技術】従来より離れた場所にいる人とコミュニケーションを図る手段として、音声情報に加え映像情報も利用したテレビ電話若しくはテレビ会議システムが知られている。

【0003】上記テレビ電話若しくはテレビ会議システムに用いるテレビカメラ装置の制御動作を音声で行う技術が既に知られており、主音声である会話より制御音声を分離する手段として、スイッチの切り替え操作で行う手法とは別に、主音声である会話より制御音声を2つのマイクを用いて分離する例が特開平5-61497号公報に開示されている。また、集音した音声を周波数分析を行うことで主音声である会話より制御音声を分離する例が特開平5-173592号公報に開示されている。

## 【0004】

【発明が解決しようとする課題】しかしながら、通信ネットワークの進歩に伴い、映像と音声を交えて不特定多数の相手と通信を行う機会が可能となった結果、相手に公開したくない映像、音声等自分のプライバシーを守る必要が生じているが、従来のキーボード、マウス等によるテレビカメラ装置の制御は操作が繁雑であり、また、主音声である会話より自動的に制御音声を分離する手法においては、機器の構成が極めて複雑であった。

【0005】本発明は、このような事情に鑑みて創案されたものであって、接続するコンピュータ装置の種類に依存することなく、手の動きのみで制御モードに移行することで、相手に公開したくない会話音声の送出遮断を容易に実現すると共に、制御モードにおいては主音声である会話に影響されずにあらかじめ定められた特定語彙のみの制御音声の認識のみを行うので、音声認識機構の簡易化が図れ、相手への送出音声の加工を音声制御で容易に実現することができるテレビカメラ装置を提供することにある。

## 【0006】

【課題を解決するための手段】本発明は、上記の課題を解決するために、通信回線を用いて映像を伝送するシステムに接続されるものであって、映像を撮像する撮像手段と、マイクで音声を集音する集音手段とを備えたテレビカメラ装置において、次の構成を採用している。

【0007】すなわち、請求項1記載の発明に係るテレビカメラ装置は、カラー映像の色再現性を保証するためのホワイトバランス制御を行うために常に測定しているCCD撮像素子からの電気信号レベルの履歴を保持する機能を制御部に備えることにより、その履歴よりCCD撮像素子からの電気信号レベルが急激に減少した期間が所定の時間持続したという情報を得ることで、撮像レンズ前面を手により遮光されたことを知ることができる。

【0008】また、請求項2記載の発明に係るテレビカメラ装置は、集音手段で得られる音声情報の音声認識手段と、テレビカメラ装置の外部へ送出される音声に所定の処理を行う音声処理手段と、計時記憶手段とを備える



ことにより、制御モードに於いて、マイクで集音した制御音声の外部への送出を停止すると共に、制御モードに移行した後の時間経過が計時記憶手段の設定情報に一致することで制御モードの終了を検知し自動的に通常モード動作にテレビカメラ装置を復帰させることができる。

【0009】そして、請求項3記載の発明に係るテレビカメラ装置は、請求項2記載の構成において、マイクで集音したあらかじめ定められた特定の語彙で構成される制御音声の音声認識手段によりテレビカメラ装置の動作を制御する制御手段を有してなるものである。

【0010】そしてまた、請求項4記載の発明に係るテレビカメラ装置は、制御部よりの制御で任意に点灯または消灯状態を表示できる表示手段を備えてなり、したがって、マイクで集音された音声の相手への送出有無、制御音声の認識可否、制御モードの残時間に応じた表示動作が行われるため、利用者がテレビカメラ装置がどのような動作状態にあるかを知ることができる。

【0011】

【発明の実施の形態】以下、本発明に係わるテレビカメラ装置の実施の形態について、図面に基いて詳細に説明する。図1は本発明のテレビカメラ装置を実施したテレビ会議システムの概要を示すブロック図である。

【0012】図1において、一方のテレビカメラ装置1aで撮像された映像及び音声は、映像信号線及び公衆電話音声信号線で接続されているコンピュータ装置2aに伝送された後、回線等の通信網3を経由して離れた場所にあるコンピュータ装置2bの表示部及び拡声部に伝えられる。

【0013】同様に、他方のテレビカメラ装置1bで撮像された映像及び音声は、映像信号線及び音声信号線で接続されているコンピュータ装置2bに伝送された後、公衆電話回線等の通信網3を経由して離れた場所にあるコンピュータ装置2aの表示部及び拡声部に伝えられる。

【0014】図2は図1のテレビカメラ装置1a、1bの全体構成を示すブロック図であり、また、図3は図1のテレビカメラ装置1a、1bの外観を示す斜視図である。

【0015】このテレビカメラ装置1a、1bは、装置全体の制御を行う制御部11、被写体の映像を撮像してテレビジョン映像信号に変換して出力する映像撮像部12、タイマ動作の設定値を記憶する計時記憶部14、集音された音声の中より特定の語彙を制御符号に変換する音声認識処理部15、マイクよりの音声信号を遮断したり、代わりに保留音楽を送出する等音声信号に処理を行う音声処理部16を備える。

【0016】そして、テレビカメラ装置1a、1bの前面には、撮像用のレンズ111が設けられるとともに、その横に音声を集音するマイク19と制御部11で点灯制御されるLED13がケース24より露出して配置さ

れている。また、テレビカメラ装置の側部には、電源スイッチ20が配置されている。

【0017】また、制御部11は、内部タイマと不揮発性メモリとを有しており（いずれも図示省略）、内部タイマは発振回路の周波数を計数することで常に正確な時刻を計時する一方、不揮発性メモリには、音声処理部16のD/A変換回路で変換することで音楽として認識可能な音符のデジタル符号、及び音声認識処理部15で制御符号として認識すべき特定の語彙が予め記憶されている。

【0018】さらに、撮像部12は、光を電気信号に変換するCCD撮像素子113、被写体の像をCCD撮像素子113の表面に光学的に結像させるためのレンズ111、CCD撮像素子からの電気信号を1秒間に60枚の映像で構成するNTSCテレビジョン映像信号または1秒間に50枚の映像で構成するPALテレビジョン映像信号に変換する信号処理回路112、テレビジョン映像信号に変換するための同期関係の同期信号発生器115、CCD撮像素子を駆動するためのタイミングパルス発生回路116及び垂直ドライバ回路114で構成されている。

【0019】そして、前述の撮像部12が特許請求の範囲の撮像手段に、映像処理部13が特許請求の範囲の映像処理手段に、音声処理部16が特許請求の範囲の音声処理手段に、音声認識処理部15が特許請求の範囲の音声認識手段に、制御部11とLED13とが特許請求の範囲の表示手段に、それぞれ対応している。

【0020】次に、上記構成のテレビカメラ装置1a、1bの動作について、図4に示す状態遷移図を参照して説明する。例えば、一方のテレビカメラ装置1aの電源スイッチ部20を操作して電源を投入すると、テレビカメラ装置1aは通常動作状態にあり、映像撮像部12よりのテレビジョン映像信号は、映像出力端子17からコンピュータ装置2aに伝送される。

【0021】同様にマイク19で集音された会話音声は制御部11により設定される増幅度に基づいて音声処理部16で音量調整された後、音声出力端子18からコンピュータ装置2aに伝送される。また、テレビカメラ装置1aの表示素子であるLED13は常に点灯している状態にあり、通常動作モードでテレビカメラ装置が動作していることを利用者に報知している。したがって、コンピュータ装置2aの表示部には、被写体の映像が表示されると共に、会話音声スピーカーより聞こえている状態にある。

【0022】この通常動作モードにおいて、コンピュータ装置2b側の利用者に聞かれたくない会話を行いたい事由が発生した場合、テレビカメラ装置1aの利用者は、従来の電話の送話口を手で覆う習慣を変えることなく、テレビカメラ装置1aのマイク部を覆う動作を1秒以上行くと、近接して設置しているレンズ111も手で

10

20

30

40

50

覆われるために映像入力が増光状態となり、CCD撮像素子113よりの電気信号レベルが黒レベルに急激に移行した期間が1秒継続したという履歴を信号処理回路112を通じて制御部11が得るため、制御部は直ちに遮断指示を音声処理部16に行うと共に制御モードに移行する。

【0023】制御モードにおいては、制御部11より音声遮断指示が音声処理部16に与えているために、マイク19で集音された会話音声は増幅度0のため、音声出力端子18からコンピュータ装置2aに伝送することができず、当然相手であるコンピュータ装置2bにも伝わらないためプライバシーを確保することができる。また、テレビカメラ装置1aの表示素子であるLED13は常に消灯している状態に移行することで、制御動作モードでテレビカメラ装置1aが動作していることを利用者に報知している。

【0024】制御モードにおいてもマイク19による集音動作は継続されており、この集音された音声は制御モードの状態では音声認識処理部15において、音声認識動作がおこなわれているため、利用者がテレビカメラ装置の動作を変更したい場合には、誤認識を避けるために「制御」という言葉をまず発声した後、続いてあらかじめ定められ制御部11に設定された語彙からなる制御語を発声することによりその制御語毎の所定の変更がテレビカメラ装置に施される。

【0025】この時、音声認識処理部15が利用者の制御音声を認識出来た場合のみLED13が1秒周期で2回点滅動作することで、利用者に認識状況を報知する構成である。ここで、制御モードにおいて音声認識される制御音声であるが図4に示す「保留音楽」、「無音」、「小さく」、「大きく」、「X秒」、「復帰」の6種類に限定することで音声認識処理部15の誤認識率を下げ、確実な認識を可能としている。

【0026】以下この制御語毎のテレビカメラ装置の制御内容を説明する。まず、制御語として「保留音楽」を認識した場合テレビカメラ装置の制御部11は、内部に備えた不揮発性メモリにあらかじめ格納された音符のデジタル符号を、音声処理部16のD/A変換回路で変換することで保留音楽を生成し、制御モードの間は常に音声出力端子18よりその保留音楽を送出する。

【0027】また、制御語として「無音」を認識した場合テレビカメラ装置の制御部11は、音声処理部16より保留音楽が送出されている場合には、その送出を停止するため音声出力端子18からは何も音声が出送されない無音状態となる。当然制御モードに移行直後で無音送出状態の場合には何も変化は生じない。

【0028】そして、制御語として「小さく」を認識した場合テレビカメラ装置の制御部11は、通常動作モードに復帰後においてマイク19で集音した会話音声を音声処理部16で増幅して音声出力端子18より送出する

際の増幅度を従来の設定に対して小さくなる様に設定を変更する。

【0029】そしてまた、制御語として「大きく」を認識した場合テレビカメラ装置の制御部11は、通常動作モードに復帰後においてマイク19で集音した会話音声を音声処理部16で増幅して音声出力端子18より送出する際の増幅度を従来の設定に対して大きくなる様に設定を変更する。

【0030】さらに、制御語として「X秒」を認識した場合テレビカメラ装置の制御部11は、通常動作モードへ自動復帰する迄の設定時間を記憶しているの計時記憶部14の設定時間を、制御語として例えば10秒と認識した場合は10秒に変更する。

【0031】さらにまた、制御語として「復帰」を認識した場合テレビカメラ装置の制御部11は、内部タイマで計時している通常動作モードより移行した後の経過時間が、計時記憶部14の設定時間に達していない状態であっても直ちにLED点滅モードに移行する。

【0032】制御語毎のテレビカメラ装置の制御内容の説明は以上であるが、制御動作モードにおいては、通常動作モードより移行した後の経過時間が制御部11の内部タイマで常に計時されており、計時記憶部14の設定時間と一致することでLED点滅モードに移行する。

【0033】LED点滅モードにおいてはLED13が1秒周期で5秒間点滅動作を行い利用者に通常動作モードへ移行することが近いことを報知するとともに、5秒間経過後は音声出力遮断が解除され、マイク19で集音された会話音声は制御部11により設定される増幅度に基づいて音声処理部16で音量調整された後、音声出力端子18からコンピュータ装置2aに伝送される通常動作モードに復帰する。

【0034】通常動作モードに復帰するとテレビカメラ装置1aの表示素子であるLED13は常に点灯することで、通常動作モードでテレビカメラ装置が動作していることを利用者に報知する。したがって、コンピュータ装置2aの表示部には、被写体の映像が表示されると共に、会話音声もスピーカーより聞こえている状態に復帰する。

【0035】以降、レンズ111が1秒間以上遮光される毎に制御モードと、通常動作モードへの復帰動作は繰り返されるが、当然のことながら制御モード状態で電源スイッチ20を遮断したとしてもテレビカメラ装置への通電は遮断され制御部がリセットされるため、再度電源スイッチ20を投入した際は、通常動作モードで動作することとなる。

【0036】

【発明の効果】本発明のテレビカメラ装置は、上記のような構成であるから、請求項1記載の発明は、相手に知られたくない事情が生じてテレビカメラ装置で現在の会話状況が集音されることを一時的に避けたい場合には、

7

送話の停止あるいは保留音楽を会話音声に変えてコンピュータ装置等を通じて伝送できるため、プライバシーを確保することができる。

【0037】また、請求項2記載の発明は、テレビカメラ装置のレンズ部を遮光するという簡単な操作で容易に移行が可能で、且つ自動的に復帰動作を行う制御モードを設けることで容易にテレビカメラ装置の動作状態の制御手段を提供することができる。

【0038】そして、請求項3記載の発明は、制御モードにおいて集音される音声のみに限定して制御符号として認識すべき制御音声の音声認識を行うことで、制御語彙の場合は会話音声と識別が容易になるような言葉を前置する等、会話音声に左右されることなくテレビカメラ装置の制御動作を音声により行うシステムを容易に構成することができる。

【0039】そしてまた、請求項4記載の発明は、マイクで集音された音声は相手に送出されているかどうかの確認が容易なだけでなく、制御モードでの制御音声の認識状況、制御モードより通常モードへの復帰が迫ってい

8

ることを視覚的に容易に確認可能とすることができる。

【図面の簡単な説明】

【図1】本発明のテレビカメラ装置を実施したテレビ会議システムの概要を示すブロック図である。

【図2】図1のシステムを構成する場合に使用される本発明のテレビカメラ装置の一実施の形態の構成を示すブロック図である。

【図3】本発明のテレビカメラ装置の外観を示す斜視図である。

10 【図4】本発明のテレビカメラ装置の動作説明に供する遷移状態図である。

【符号の説明】

1a, 1b テレビカメラ装置

11 制御部

12 映像撮像部

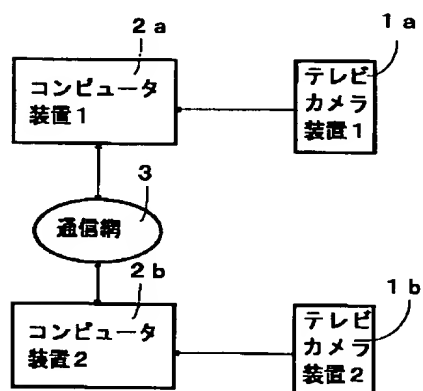
14 計時記憶部

15 音声認識処理部

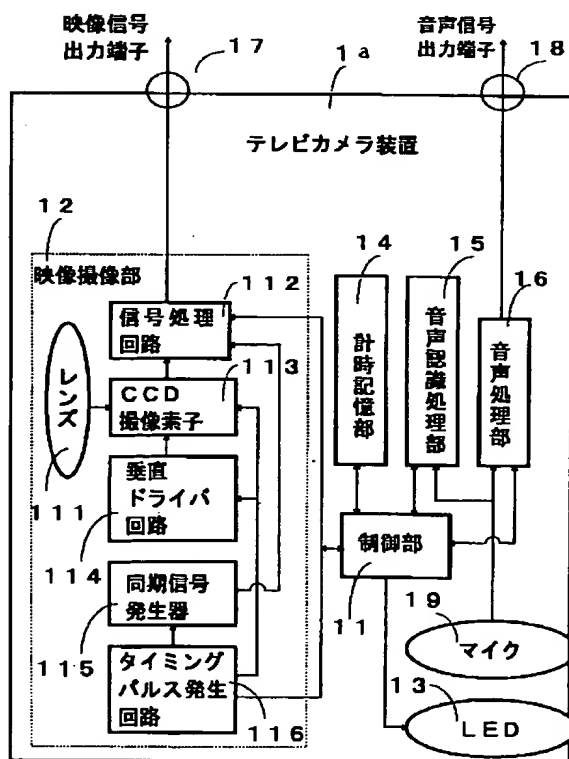
16 音声処理部

20 電源スイッチ部

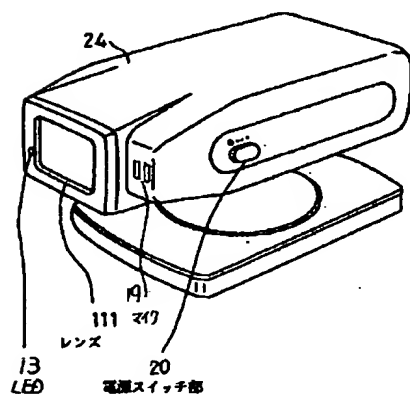
【図1】



【図2】



【図3】



【図4】

